人工智能领域高产科研团队的演化研究*

■ 邹本涛[1] 王曰芬[1,2] 余厚强[1,2]

[1] 南京理工大学经济管理学院 南京 210094 [2]江苏省社会公共安全科技协同创新中心 南京 210094

摘 要：[目的/意义]团队合作已成为当今知识创新的一种重要组织形式。从动态网络视角探究科研团队动态演化规律对于促进科研团队的发现、组建和管理具有重要意义。[方法/过程]以人工智能领域为例，采用Louvain社群发现算法识别人工智能领域研究团队，通过计算团队合作网络中节点数、边数、网络密度和平均聚集系数四项拓扑指标的极值分布，从微观和宏观视角探究该领域高产团队演化的特征与规律，以揭示科研团队演化的内在动因。[结果/结论]微观视角下，合著网络拓扑指标的极值分布揭示人工智能领域高产团队演化的动态属性；宏观视角下，高产团队在网络密度与网络平均集聚系数上呈现出演化共性，多数团队在演化中催生更多新的合作关系的产生；结合团队的演化路径来看，人工智能领域高产团队中各时期的"小团体"合作现象显著，且"小团体"之间的合作直接影响着整体团队的走向。

关键词：人工智能 研究团队 极值分布 动态演化 拓扑结构
分类号：G250
DOI：10.13266/j.issn.0252-3116.2020.20.003

1 引言

科学研究问题的跨学科性和复杂性在日益增加，推动着科研团队汇集更多不同领域的专家开展协作式科研工作[1]。值得提出的是，这种团队式协作在科学诞生之初就成为科研活动的内在属性，近几十年的研究表明其仍旧是当今科学研究的重要特征之一[2-3]。并且，科研团队在促进科学发现与产出方面有着重要贡献[4]，比如稳固的团队更有可能带给其成员更多的JCR(Journal Citation Reports)期刊论文产出[5]、有机会承担更多研究项目和申请到更多专利[6]。通过对科研团队进行发掘与分析，不仅能够厘清特定领域发展的内在驱动因素[7]，还能引导处于萌芽期或发展期的团队顺利过渡到成熟与稳定的阶段，进一步提高科研产出与推动科学进步。因此，众多学者对来自计算机科学、社会学、物理学和生物学等领域的团队进行不同视角下的探索分析[8]，甚至已形成"团队科学"(Science of Team Science)这一特定研究领域，重点关注科研团队的组织、交流和研究开展的过程[9]。

早期的研究工作主要从静态网络视角出发，以某一时间段或某一时刻的团队为研究对象，将其看作是恒定不变的、静态的结构实体[10-11]。但在实际情况中，这种存在交互行为的网络是随时间而不断变化的[12-13]。团队中的成员以及成员之间的关系是动态变化的，比如新成员加入某个团队后，会评判该团队是否能实现其个体层面的目标，而团队则会评估该成员是否能提升团队的整体价值，这二者之间的匹配程度将决定该成员能否和团队维持稳定的关系[14]。若将研究对象限定为科研团队，其演化将受到更多因素的影响，如许多科研团队的形成是出于申请或完成特定课题这一目的临时"拼凑"而成[15]，科研合作还受学者间的地理距离和其科学知名度、可见度、认可度等因素影响[16]。因此，研究者不再仅分析团队的静态网络属性，而是更进一步地去分析团队的演化并试图揭示影响演化的因素[17]。现有的研究主要集中在：①通过识别事件来表征与发现演化状态[21]，如Y. Wang等[22]定义诞生(birth)、分裂(split)、融合(merge)和消亡(death)四种事件，从网络规模和网络成员等方面探讨社区的动态演化特征；G. Palla等[24]识别诞生(born)、成长(grow)、分裂(split)、融合(merge)、收缩(contract)

* 本文系国家社会科学基金重大项目"面向知识创新服务的数据科学理论与方法研究"(项目编号:16ZDA224)研究成果之一。
作者简介：邹本涛(ORCID：0000-0002-3972-0705)，博士研究生；王曰芬(ORCID：0000-0002-7143-7766)，教授，博士生导师，通讯作者，E-mail：yfwang@njust.edu.cn；余厚强(ORCID：0000-0002-9241-6630)，副教授，博士。
收稿日期：2020-05-11 修回日期：2020-07-06 本文起止页码：23-33 本文责任编辑：徐健

和消失(disappear)等事件,从平稳性(stationarity)角度分析其与社区年龄和规模之间的相关性;②通过拓扑结构变化来表征与反映演化的整体面貌[19],如 A. L. Barabási 等[20]计算滑动时间窗口内合著网络拓扑指标的演变,指出演化过程中作者倾向于选择拥有较多合著者的学者进行合作;郭世月等[27]则分析团队在规模、网络密度和平均距离等拓扑指标上的演变;③结合主题等语义信息通过主题关联分析演化过程,F. Osborne 等[25]指出应关注作者和研究主题之间的关系随时间的演化;L. Backstrom 等[28]提出探讨团队的演变是"作者驱动"(Topics follow people)还是"主题驱动"(People follow topics);④通过概念外延将研究社区当成团队进行演化分析[16,26],以挖掘出网络结构属性的转变[18]、演变特征[12]以及其发生动态变化的内在驱动力[7]。

目前,应用于团队演化分析的方法主要有两类(国外研究多以"社区"这一概念指代包含"团队"在内的多种实体):其一是先将数据集划分成多个时间片段,识别各片段中的社区后再比较不同片段间社区的相似度,从而获得社区的演化路径;其二是演化社区发掘方法,根据前一个时间片段中识别的社区结果,结合节点、边的相关属性来挖掘下一阶段的演化[21]。尚未有研究证实哪种方法更加有效,从适用性角度来说,第一类方法能较好应对不稳定变化社区。

综上所述,深入探索科研团队的演化特点与规律,探究有效可行的分析方法,不仅具有重要的学术价值,而且对于支撑政策制定具有实践意义。在已有的研究中,从网络拓扑角度进行的分析视角较为单一,相关研究主要针对所有识别的团队,从整体视角分析团队或整体合著网络中拓扑指标随时间的变化情况。虽然这种做法可以揭示部分团队的演化,但是不同团队对领域发展所作出的贡献是不同的,使得"部分团队"无法针对性地反映出不同团队的发展。同时,越来越多的研究表明,研究团队在演化过程中具有较高的不稳定性[22-23]。因此,本文以人工智能领域的高产研究团队为研究对象,结合团队演化路径从微观与宏观双重视角切入,通过计算团队合作网络中节点数、边数、网络密度和平均聚集系数拓扑指标的极值分布,对高产团队拓扑结构的演化进行分析,以期在一定程度上解释高产团队的形成与内在演变原因,并促进学术研究与实践应用的发展。

2 研究设计与团队选取

基于本文研究目的及上述梳理,从数据分析出发,按照如图 1 所示研究设计图开展本文工作。

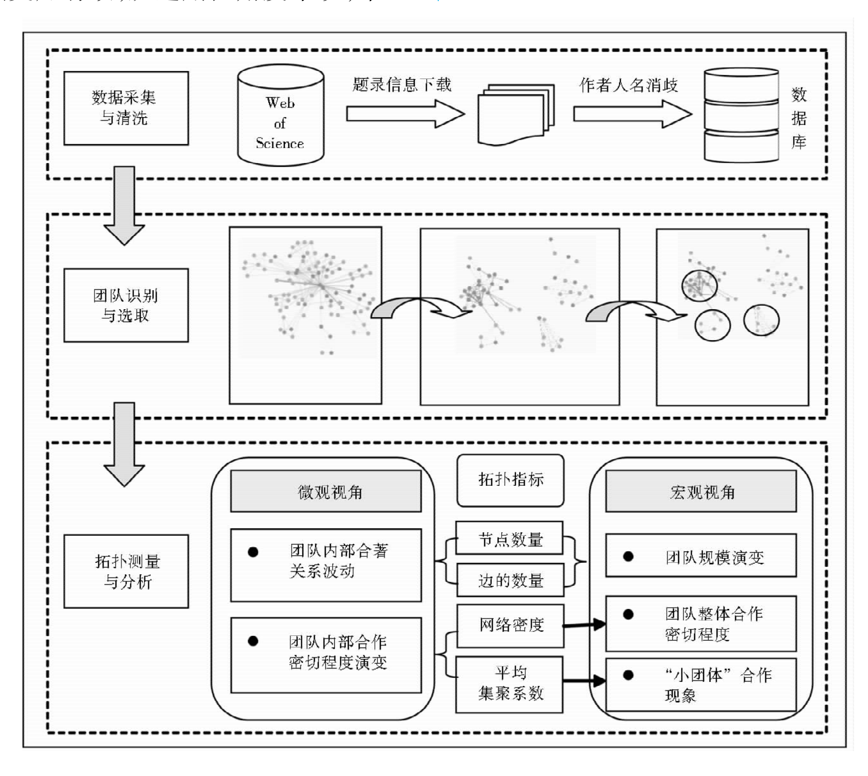


图 1 研究设计图

邹本涛，王曰芬，余厚强.人工智能领域高产科研团队的演化研究[J].图书情报工作,2020,64(20):23－33.

首先,在数据采集与清洗环节,本文选取 Web of Science 中核心合集为数据源,设定检索式为 WC ＝ “Computer Science, Artificial Intelligence”,时间限定为 2009－2018 年,于 2019 年 1 月检索到共计 42 1148 条记录。作者人名歧义问题严重影响合著网络的构建,而合著网络的质量直接决定团队识别结果的准确性。因此,笔者首先结合合著者及作者机构两类信息完成人名消歧工作,将清洗后的数据导入 MySQL 数据库中存储。

然后,进入到科研团队的识别与选取阶段。在构建好整体的作者合著网络之后,笔者选用 Louvain 社区探测算法来识别人工智能领域的科研团队。这一步识别出科研团队 23 423 个,共涉及作者 656 668 名。发文量一直作为传统的计量指标被广泛使用,但针对高发文量团队进行动态演化分析的研究还比较少见,本文将发文量位列前 100 名的团队界定为高产科研团队,并将此部分团队作为研究对象开展后续分析。

最终,进行拓扑测量与演化分析。在获取待分析团队之后,匹配团队成员发表的文章集合,将文章去重后按照发表年份(2009－2018 年)划分成 10 个时间区间。在每个区间内,获取文章对应的所有作者,再构建特定团队在该区间内的合著网络,各个网络按时间串联起来便形成团队的演化路径。针对每一个时间区间内的合著网络,笔者计算该网络的节点数量、边的数量、网络密度和网络平均集聚系数四项拓扑指标。团队拓扑结构的测度是从网络分析角度来理解科研团队的结构特征,而对每个时间片段中合著网络拓扑结构的分析,将揭示团队演化过程中网络结构的变化。与此同时,还可以挖掘高产科研团队演化过程中展现的共性和特性,以加深对团队动态演化的理解。基于此,本文结合微观与宏观双重视角对科研团队的动态演化进行分析。在微观视角下,本文探索所选拓扑指标极值的年份分布情况。其中,点和边数量的极值分布直接反映成员与成员关系(合著关系)的波动情况,网络密度和平均集聚系数反映团队演化过程中成员的合作密切程度,其极值分布可以在一定程度上反映高产团队内部合作关系的波动。在宏观视角下,本文分析拓扑指标在十年跨度上的整体演变。其中,点数量的变动反映团队规模的变化,边数量的变动反映团队成员合著关系的变化,网络密度的演变代表着团队整体合作的密切程度,平均集聚系数的变化则更加具体的体现出“小团体”合作现象的演变。

3 基于微观视角的高产科研团队拓扑结构演化及其分析

3.1 点和边数量的极值分布

拓扑指标极值的年份分布可以直观揭示对应指标的波动情况,这种波动的内在原因则是团队演化的具体表现。图 2 和图 3 分别是所选团队在演化过程中点和边数量极值的分布。两个图反映二者呈现出高度的一致性:人工智能领域多数团队点和边数量的最小值分布在前五年和最后一年,最大值分布在所选研究区间的中后期(2012－2017 年)。这种分布说明团队在最初几年里,点和边的数量相对较少,但是随时间变化,各团队在这两方面逐渐呈现不同的演化行为——有的团队在中期达到最大值,说明团队中点和边的数量在后期会经历减少的过程;有的团队则在后期达到最大值,意味着团队中点和边数量或持续增加、或“增中有减”。

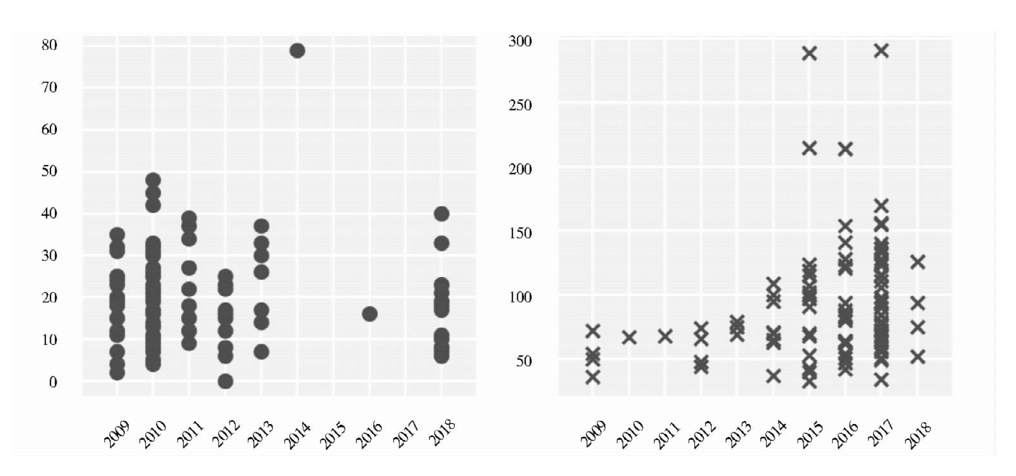


图 2 节点数的最小值分布(左)、最大值分布(右)(纵轴:节点数量,横轴:年份)

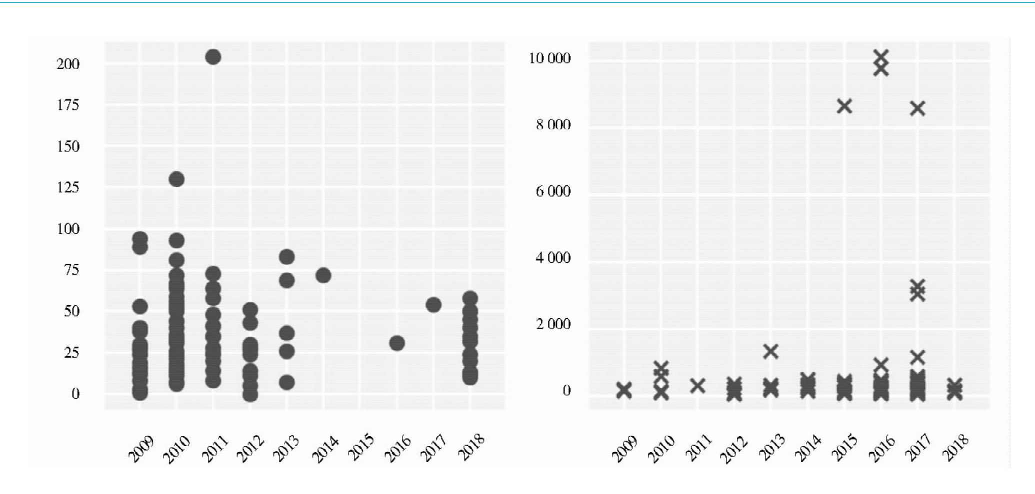


图 3 边数的最小值分布(左)、最大值分布(右)(纵轴:边数量,横轴:年份)

表 1 给出更具体的数值展示,2010 年有 33 个高产团队在节点数量和边数量上取最小值,其中有 29 个团队在这一年同时达到边和节点数量的最小值。可以看出,进行过滤之后识别出的高产团队,在节点和边的极值分布上高度同步,结合图 2 与图 3,可知其最小值主要分布于 2009 – 2012 年以及 2018 年。进一步的,表 2 列举了这些年份的示例团队(每一年选取发文量最多的 2 个团队展示),节点最小值分布在 2009 年的团队排名最高达到第 5,分布在 2018 年的团队发文量排名最高为第 17。可以发现,无论是从平均排名、最高排名或者团队的平均发文量来看,这些团队之间没有表现出较大差异。

表 1 团队节点、边取最小值时的年代分布

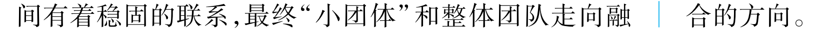
年份	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
节点	19	33	11	11	7	1	0	1	0	17
边	18	33	15	11	5	1	0	1	1	15
相同团队	14	29	10	8	5	0	0	1	0	13

表 2 团队示例:节点最小值主要分布年份

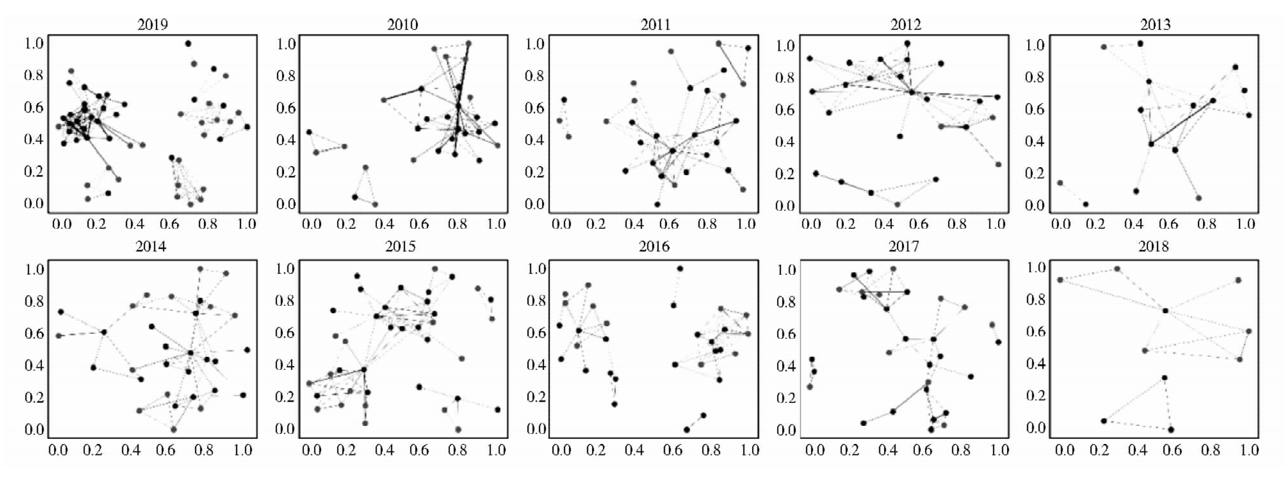
团队编号	发文量(篇)	代表成员(发文量前 3)	成员数(位)	排名/平均排名	最小值所在年份	平均发文量(篇)
342	207.743 5	Xu, Zeshui_2;Zhai, Yuling;Liao, Huchang	49	5/46	2009	153.368 3
1800	197.665 3	Zafeiriou, Stefanos;Panagakis, Yannis_2;Pantic, Maja	48	8/46	2009	
448	242.960 1	Aliev, Rafik A.;Pedrycz, Witold_2;Dlugosz, Rafal	52	1/55	2010	147.739 7
1 927	219.168 5	Browne, Will N.;Kukenys, Ignas_2;Zhang, Mengjie_2	54	2/55	2010	
207	189.131 5	Jiao, Licheng;Liu, Hongying_4;Yang, Shuyuan_3	98	12/52	2011	148.042 6
276	172.285 9	Yaqoob, Naveed;Yousafzai, Faisal_3;Zeb, Anwar	61	19/52	2011	
1064	213.490 3	Melin, Patricia;Mendoza, Olivia;Castillo, Oscar	58	3/34	2012	166.291 0
203	205.230 1	Xian, Yongjin;Zheng, Cheng-De;Wang, Yingchun_3	73	6/34	2012	
736	173.503 4	Hoogendoorn, Mark;van Maanen, Peter-Paul_1;Treur, Jan	52	17/61	2018	140.576 8
1 326	173.165 9	Tsai, Yu-Chuan;Wang, Shyue-Liang;Kao, Hung-Yu	50	18/61	2018	

注:发文量采用分数计数法计算;姓名后编号用于区分同名作者,下同

结合演化周期来看,不同团队处于不同的"生命阶段",造成团队间在拓扑指标上呈现不同的分布。其中,在 2018 年达到节点和边最小值的团队,其成员数量在减少,成员间合著强度也在减弱,表现在团队层面则对应于团队的"衰落",图 4 即为代表性团队的演化过程。这些团队的演化特征是小团体合作现象比较明显,在多个演化窗口中均可以看到其形成紧密联系的"小团体"。最终,"小团体"之间的平衡作用没有使整体团队走向汇合,导致整体团队在后期的"衰落"。与之相对应的另一种演化模式以编号为 207 的团队为代表,该团队以西安电子科技大学焦李成作为带头人。带头人在 10 个时间窗口中均有出现,演化的前期、中期与后期均能看到以其为核心的合著网络。同时,虽然也有"小团体"合作现象,但"小团体"与整体团队之


邹本涛，王曰芬，余厚强.人工智能领域高产科研团队的演化研究[J].图书情报工作,2020,64(20):23－33.

间有着稳固的联系，最终“小团体”和整体团队走向融合的方向。



注：横轴与纵轴范围均为0－1，为网络布局标准化设置，无实意，下同

图4　团队#736的演化路径

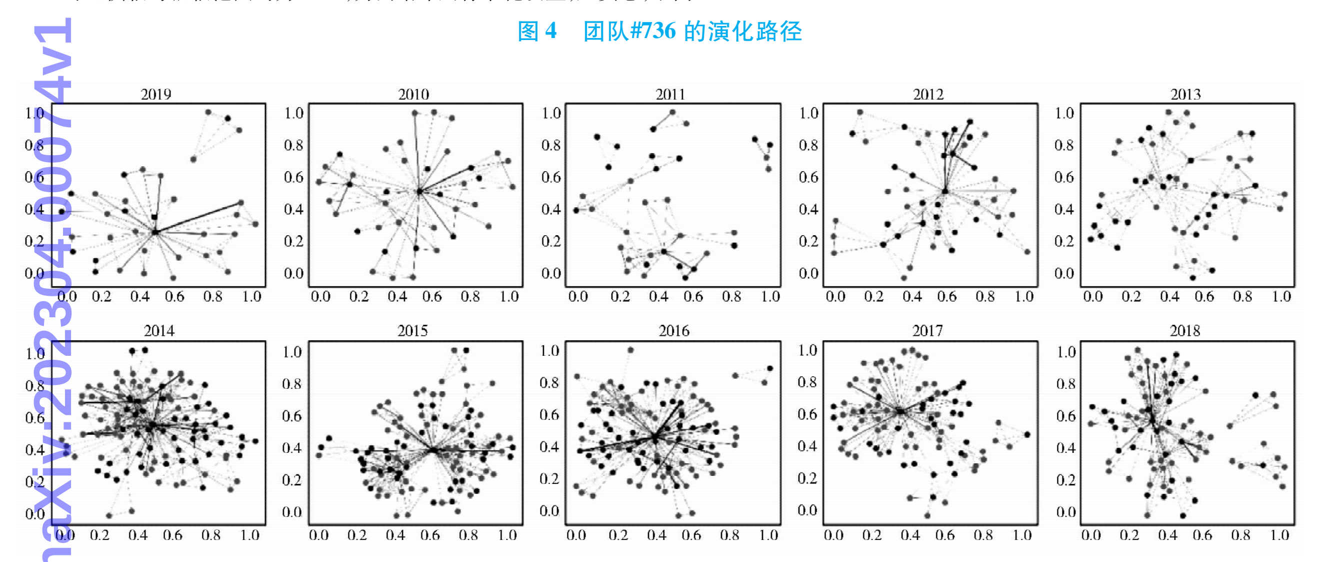


图5　团队#207的演化路径

3.2　网络密度和网络平均集聚系数的极值分布

与点和边数量的极值分布相反，团队演化过程中网络密度的最小值分布集中在所选研究区间的中后期阶段(2014－2017年)，但最大值主要分布在前面5年和最后一年。结合图2和图3的分布来看，一些团队在最初几年的成员较少，但团队内部之间的合著较多，于是呈现出网络密度的最大值分布集中在前几年，如图6所示。图7表示团队的平均集聚系数分布，最大值和最小值的分布相对前三种指标更加均匀，但主要分布于前5年。平均集聚系数最小值大多数介于0.4－0.85之间，最大值主要介于0.8－1.0之间，反映出团队普遍有较高的集聚程度。同时，平均集聚系数的均匀分布也在一定程度上反映存在较多小团体合作现象。

结合社会网络的理论来看，网络密度反映着科研团队中合作关系的疏密。对于处于紧密合作关系中的高产团队，后期的合作关系将趋于稳定(图6左所示)，此时的团队中增加一个成员，可能实际只建立几条边的合作关系，但潜在合作关系增加数十倍，于是团队的网络密度会在后期更可能取到最小值。对于网络平均集聚系数，“三人组”小团体合作现象是最直观的反映。当合作关系发展到一定程度时，“闭合三元组”的数量更多，也就使得一些团队的网络平均集聚系数在后期取得最大值。而高产团队的高发文量，意味着团队本身在发展过程中有部分比较稳定的合作关系，于是在中前期也有一些团队取得网络平均集聚系数的最大值(图7右所示)。

同样的，仍以图4和图5所选团队为例：图4所示#736团队在2018年取得网络密度的最大值，在2017年取得最小值。该团队在2018年节点数量急剧减少，导致网络密度增加；而在2017年，网络节点数处于中

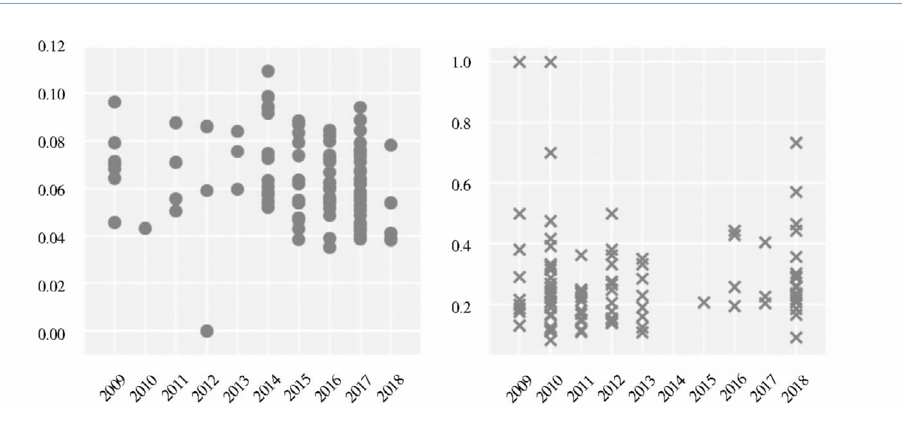


图 6　网络密度最小值(左)、最大值(右)分布(纵轴:网络密度,横轴:年份)

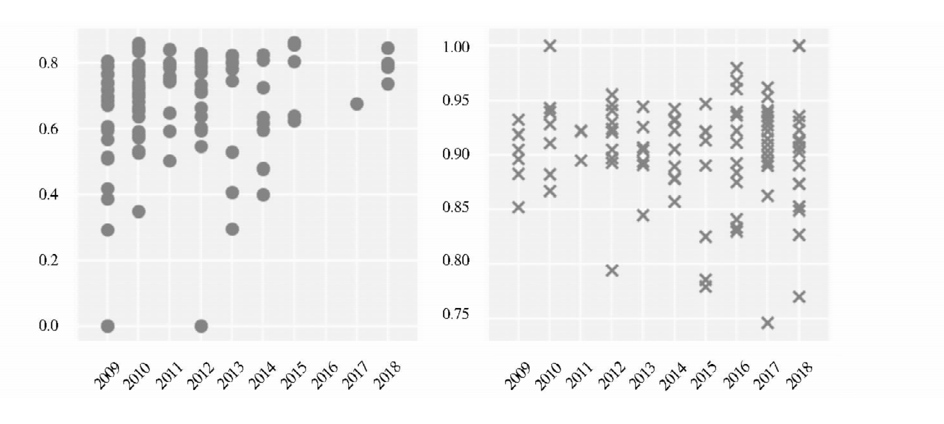


图 7　平均集聚系数最小值(左)、最大值(右)分布(纵轴:平均集聚系数,横轴:年份)

等水平,边的数量处于较低水平,反映团队成员之间的合作程度较低。对于网络平均集聚系数,其分别在 2013 年和 2018 年取得最小值和最大值。观察其演化路径图可解释这种极值分布,团队中"闭合三元组"数量在 2013 年占比较低,而在 2018 年,几乎所有的点都处于"闭合三元组"之中。图 5 所示#207 团队分别在 2010 年和 2014 年取得网络密度的最小值和最大值。从演化图来看,该团队在 2010 年以焦李成为中心呈辐射状,但辐射出去的点之间没有紧密联系,从而有最小网络密度;而在 2014 年,网络呈现以焦李成、公茂果和马文萍为中心的辐射网络,且辐射边缘的点之间有较多的边相连,带来网络密度最大值。对于网络平均集聚系数,其最小值和最大值分别在 2012 年和 2011 年取得,分别为 0.819 7 和 0.922 2,由于团队内部合作一直较为紧密,"闭合三元组"的占比稳定在较高水平,使得合著网络在演化中集聚程度较高。

4　基于宏观视角的高产科研团队拓扑结构演化及其分析

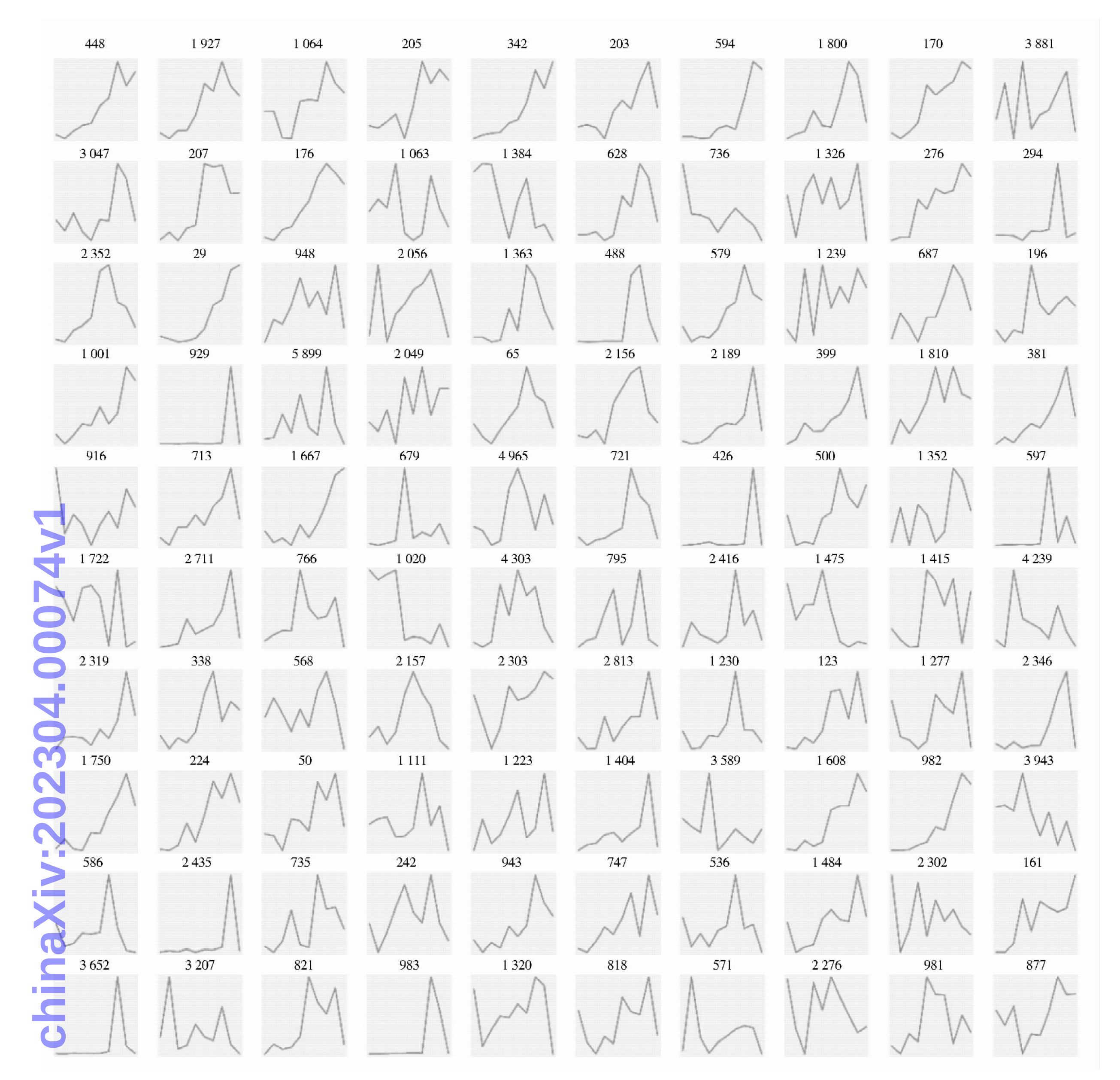
4.1　点和边数量的整体演变

笔者以每个团队在 4 个指标上的演化为基础,试图挖掘这些高产团队在某种指标或某几种指标上演化的共性,以辅助解释团队的动态演化。具体的,通过测度团队在每个时间窗口中的拓扑指标,可以画出各指标随时间的变化。图 8 是所有高产团队边的数量随时间的变化,从其所展示的演化来看,各个团队没有表现出明显的共性特征,但可以得知多数团队边的数量呈增加趋势,反映出新的合作关系在演化中逐渐增加。节点数量的演化与之类似,受篇幅限制此处不对节点数量演化做进一步讨论。

4.2　网络密度和网络平均集聚系数的整体演变

值得注意的是,团队演化过程中平均集聚系数和网络密度的变化有两种明显的演化类型,图 9 反映的是 5 个团队的示例。两种演化类型具体表现为:类型一中团队的平均集聚系数处于中高水平并保持稳定,图 8 中团队编号为 1 064、205 和 203 皆属于此类;类型二有 1 – 2 年间的平均集聚系数较低,其余时间与前一种类型一致,如图 9 中团队 342 和 1 363。此外,网络密度与平均集聚系数在演化中基本表现出相反的趋势:多数团队的密度稳定在较低水平,非常少的团队在 1 – 2 年间有波动至中高水平。并且,二者的这种变化能保持较高程度的同步,当网络密度变动时,平均集聚

邹本涛，王曰芬，余厚强.人工智能领域高产科研团队的演化研究[J].图书情报工作,2020,64(20):23－33.



注：每个子图中，纵轴为边数量，横轴为年份

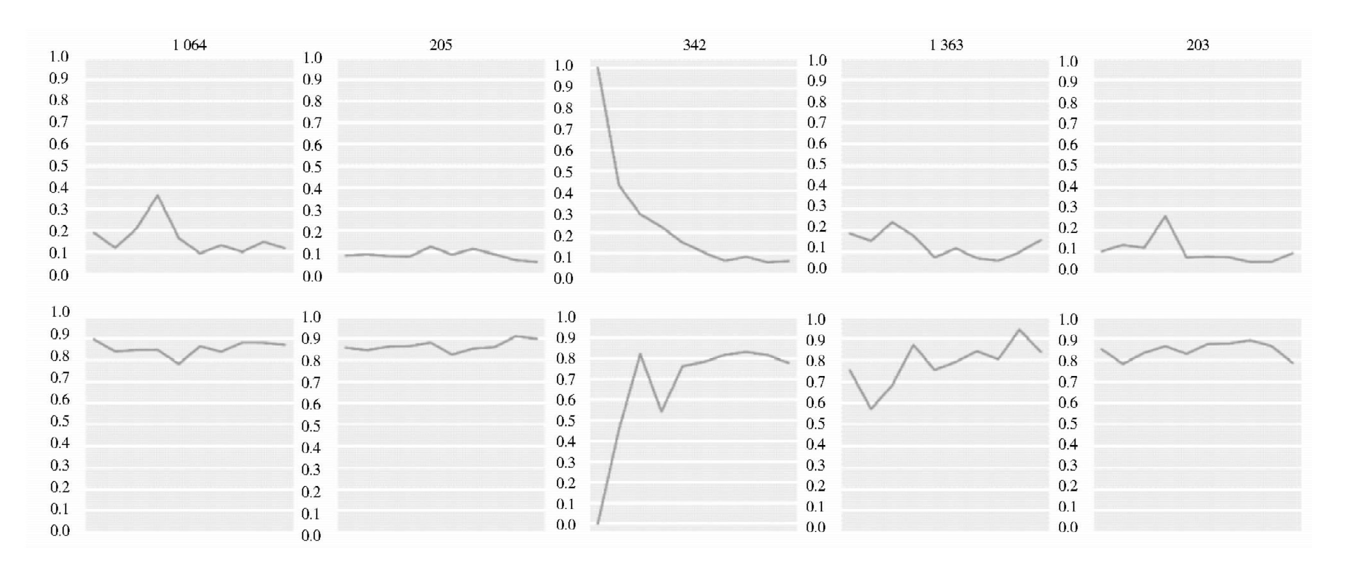
图 8　所有高产团队边数量的演化

系数在 1－2 年间会发生波动，反之亦然。此外，这两种指标发生波动的时间基本分布在最初的 5 年，与上一小节中团队节点数量和边数量极值的分布情况相对应。

图 10 和图 11 分别选择团队 205 和 342 为例来阐释这两种类型的演化过程。团队 205 在 10 个时间窗口中均表现出明显的“小团体”合作现象，但这些“小团体”之间的合作在前 5 年并不紧密，2014 年之后可以看出“小团体”之间开始发展出紧密的合作关系。并且，2014 年正对应于图 9 中团队 205 网络密度发生波动的年份。团队 342 的演化过程揭示出该团队的成长过程，前 5 年中点和边的数量都呈缓慢增长，2014 年之后形成较为稳定的内部合作关系，其后这些关系得到不断稳固，进一步发展出“小团体”现象，直到最后两年，各团体之间产生较多合作关系。

5　讨论与总结

通过上述分析发现，首先，微观视角下的分析非常直观的反映出人工智能领域高产科研团队演化的动态属性，团队中点的数量、边的数量、网络密度和网络平均集聚系数的极值分布不一。具体地，点和边数量的最小值分布在前五年与最后一年（2009－2014 年与 2018 年），点数量的最大值集中在与之相对的 2015－2017 年，边数量的最大值则较为均匀的分布在 2013－



注:纵轴为对应指标,横轴为年份

图 9 团队网络密度(上)和平均集聚系数(下)演化类型示例

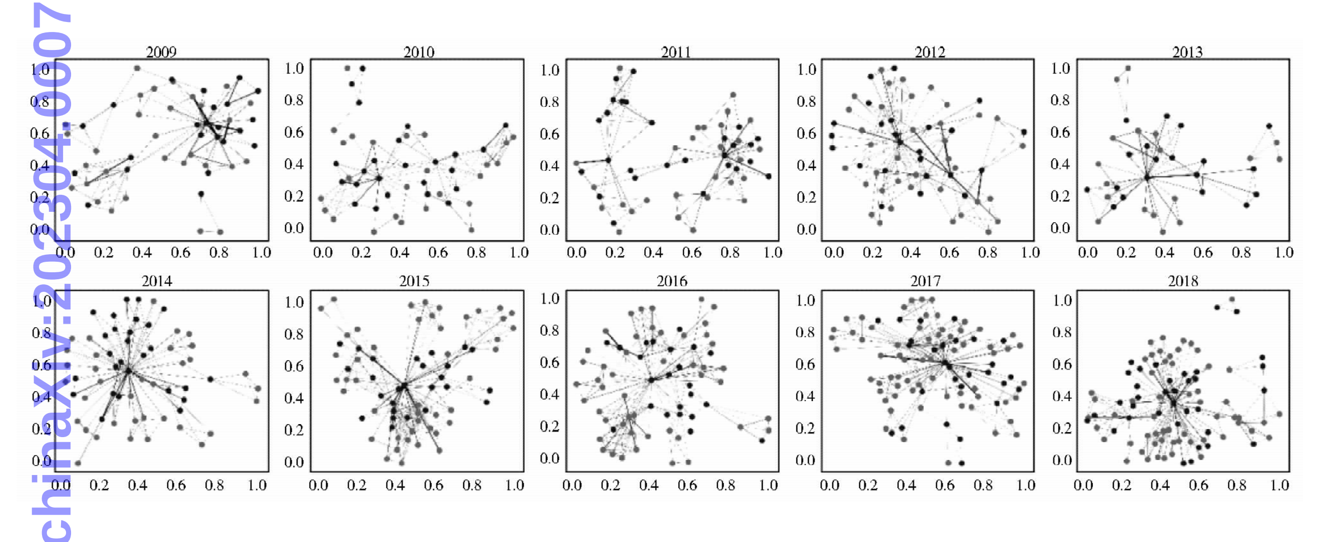


图 10 类型一团队的演化路径(以团队 205 为例)

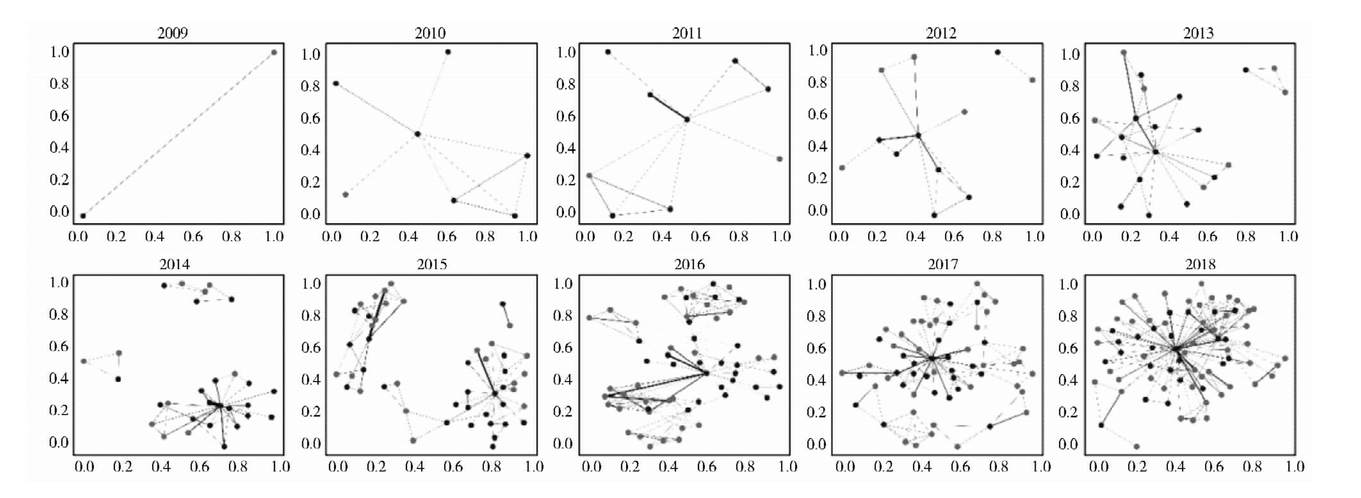


图 11 类型二团队的演化路径(以团队 342 为例)

2017年;网络密度的最小值集中分布在2014－2017年,最大值分布于2009－2014年和2018年;平均集聚系数的最小值集中分布于2009－2014年,而最大值在所有年份均有分布。对比可知,该领域高产团队的点数量、边数量和平均集聚系数最小值基本分布在同一时间段(2009－2014年),而这一时间段正好是网络密度最大值的分布区间。

其次,宏观视角下对团队拓扑指标整体演变的分析与微观视角下所作分析相印证:人工智能领域高产科研团队点和边数量的演化都呈增加趋势,部分团队在后期走向分解,对应的点和边数量再经历减少阶段;团队网络密度和平均集聚系数的演化呈现出两种模式,一种是这两项指标均处于相对稳定的水平,另一种是当某一指标发生波动时,另一指标则会在接下来的两年内发生相反的变化。综合两种分析视角和团队演化路径来看,人工智能领域大多数高产科研团队都能不断催生更多新的合作关系——平均集聚系数指标揭示出这种合作关系中有非常明显的"小团体"合作现象,且"小团体"之间的合作关系也趋于稳定。虽然,这些现象及其演化,需要进一步得到领域专家结合人工智能的领域特点来做深入解释。但是,从初步结果推断来看,人工智能自身的领域特性导致高产团队表现出两个方面的演化行为:其一,该领域以数学和计算机理论与技术为基础,解决跨学科的综合性问题,极大推动着合作行为的产生。当一个团队能成长为高产团队时,其团队中的"小团体"要么与团队带头人紧密联系,要么"小团体"之间逐渐走向联合;其二,相关技术在子领域间的应用有较大"流通性"。比如特征提取技术同时应用于计算机视觉任务和自然语言处理任务中,使得研究计算机视觉的"小团体"有较大可能性与研究自然语言处理的"小团体"建立联系。

综上,本文的主要研究结论可提炼为以下几点:

(1)从微观视角进行的分析发现,合著网络拓扑指标的极值分布揭示出人工智能领域高产科研团队演化的动态属性。该领域内,高产科研团队的合著网络中点和边数量的极值分布高度关联,并且这种关联分布与网络密度和网络平均集聚系数的极值分布互相印证,各项指标的极值分布揭示出,即便是高产团队的成员和内部关系也处于较大变化中。

(2)宏观视角下所做的分析表明,人工智能领域高产科研团队在网络密度和网络平均集聚系数的演变上表现出一定共性,且多数高产科研团队能不断催生新的合作关系(原有成员之间产生新的合作或随新成员加入带来的合作),这意味着,多数团队在成员数量和成员关系的动态演变中能保持一些拓扑结构的稳定。

(3)无论是从微观视角还是宏观视角展开分析,在任一种视角下深入观察和分析团队的演化路径时,均可发现人工智能领域高产科研团队演化过程中显著的"小团体"合作现象。并且,"小团体"之间的合作关系直接影响整体团队的未来走向,"小团体"间合作更紧密时,整体团队的网络密度和平均集聚系数均将提高;"小团体"之间疏于合作时,整个团队甚至可能走向"衰落"阶段,成员数量和成员之间的合作关系都有减少。

(4)人工智能领域高产科研团队在演化中除了表现出一定的共性规律,不同团队间的演化行为仍存在较大差异,如拓扑指标中点和边的数量增减规律的特异性,不同团队表现出不同的变动规律。

关于未来研究的展望,拟从两方面进行:首先,本研究中所分析出的团队演化共性还有待深入。比如,边的数量和节点数量呈现出增加趋势,需进一步明确边数量的增加是来自"新成员"与"旧成员"的合作还是"旧成员"之间的合作;再如,可尝试引入新的分析视角或增加高产团队数量以进一步探析演化共性;其次,所选取的研究对象是非常有针对性的高产科研团队,还可以进一步分析由其他指标遴选出的科研团队的表现(如高被引团队),以探索相关操作与结论的推广性。

参考文献:

[1] VERBREE M, HORLINGS E, GROENEWEGEN P, et al. Organizational factors influencing scholarly performance: a multivariate study of biomedical research groups[J]. Scientometrics, 2015,102(1):25－49.

[2] PERIANES-RODRÍGUEZ A, OLMEDA-GÓMEZ C, MOYA-ANEGÓN F. Detecting, identifying and visualizing research groups in co-authorship networks[J]. Scientometrics, 2010,82(2):307－319.

[3] VAN RAAN A F J. Scaling rules in the science system: influence of field-specific citation characteristics on the impact of research groups[J]. Journal of the American Society for Information Science and Technology, 2008,64(7):1852－1863.

[4] YU Q, SHAO H, DUAN Z. Research groups of oncology co-authorship network in China[J]. Scientometrics, 2011,89(2):553－567.

[5] REY-ROCHA J, GARZÓN-GARCÍA B, MARTÍN-SEMPERE M

J. Scientists' performance and consolidation of research teams in biology and biomedicine at the Spanish Council for Scientific Research[J]. Scientometrics, 2006,69(2):183-212.

[6] MARTÍN-SEMPERE M J, GARZÓN-GARCÍA B, REY-ROCHA J. Team consolidation, social integration and scientists' research performance: an empirical study in the Biology and Biomedicine field[J]. Scientometrics, 2008,76(3):457-482.

[7] MORRIS S A, GOLDSTEIN M L. Manifestation of research teams in journal literature: a growth model of papers, authors, collaboration, coauthorship, weak ties, and Lotka's law[J]. Journal of the American Society for Information Science and Technology, 2007, 64(7):1852-1863.

[8] AZAOUZI M, RHOUMA D, ROMDHANE L B. Community detection in large-scale social networks: state-of-the-art and future directions[J]. Social network analysis & mining, 2019, 9(1):23.

[9] BÖRNER K, CONTRACTOR N, FALK-KRZESINSKI H J, et al. A multi-level systems perspective for the science of team science [J]. Science translational medicine, 2010,2(49):49cm24-49cm24.

[10] GIRVAN M, NEWMAN M E J. Community structure in social and biological networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2002,99(12):7821-7826.

[11] NEWMAN M E J. Modularity and community structure in networks [J]. Proceedings of the National Academy of Sciences of the United States of America, 2006,103(23):8577-8582.

[12] ASUR S, PARTHASARATHY S, UCAR D. An event-based framework for characterizing the evolutionary behavior of interaction graphs[C]// ACM SIGKDD international conference on knowledge discovery & data mining. New York: ACM, 2007:913-921.

[13] MASMOUDI A, MEZGHANI E, BELLAAJ H, et al. A web-based knowledge management system for scientific research team[C]// 2017 IEEE 26th international conference on enabling technologies: infrastructure for collaborative enterprises. Poznan: IEEE, 2017: 296-301.

[14] OWENS D A, MANNIX E A, NEALE M A. Strategic formation of groups: issues in task performance and team member selection[J]. Research on managing groups and teams: composition, 1998, 1 (1998):149-165.

[15] JOHNSON N F, XU C, ZHAO Z, et al. Human group formation in online guilds and offline gangs driven by a common team dynamic [J]. Physical review E, 2009, 79(6):1-11.

[16] CEBALLOS H G, GARZA S E, CANTU F J. Factors influencing the formation of intra-institutional formal research groups: group prediction from collaboration, organisational, and topical networks [J]. Scientometrics, 2018,114(1):181-216.

[17] MILOJEVIC S. Principles of scientific research team formation and evolution[J]. Proceedings of the National Academy of Sciences of the United States of America, 2014,111(11):3984-3989.

[18] TAKAFFOLI M, SANGI F, FAGNAN J, et al. MODEC-modeling and detecting evolutions of communities[C]//Fifth international AAAI conference on weblogs and social media. California: The AAAI Press, 2011:626-629.

[19] 王曰芬,李冬琼,余厚强. 生命周期阶段中的科学合作网络演化及高影响力学者成长特征研究[J]. 情报学报, 2018,37(2):121-131.

[20] BARABÁSI A L, JEONG H, NÉDA Z, et al. Evolution of the social network of scientific collaborations[J]. Physica A: Statistical mechanics and its applications, 2002, 311(3-4):590-614.

[21] TAKAFFOLI M, FAGNAN J, SANGI F, et al. Tracking changes in dynamic information networks[C]// International conference on computational aspects of social networks. Piscataway: IEEE, 2011:94-101.

[22] WANG Y, WU B, DU N. Community evolution of social network: feature, algorithm and model[EB/OL]. [2020-03-31]. http://arxiv.org/abs/0804.4356.

[23] HOPCROFT J, KHAN O, KULIS B, et al. Tracking evolving communities in large linked networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101 (SUPPL. 1):5249-5253.

[24] PALLA G, BARABÁSI A L, VICSEK T. Quantifying social group evolution[J]. Nature. 2007, 446(7136):664-667.

[25] OSBORNE F, SCAVO G, MOTTA E. A hybrid semantic approach to building dynamic maps of research communities[C]//International conference on knowledge engineering and knowledge management. Cham: Springer, 2014: 356-372.

[26] 刘先红,李纲. 国家自然科学基金连续资助期间科研团队的合作稳定性分析[J]. 中国科学基金, 2016,30(4):372-378.

[27] 郭世月,李睿. 合著网络模式与论文产出数量及质量的相关性研究——以中国催化剂领域期刊论文合著网络为例[J]. 情报学报. 2013,32(3):314-327.

[28] BACKSTROM L, HUTTENLOCHER D, KLEIUNBERG J. Group formation in large social networks: membership, growth, and evolution[C]// Proceedings of the 12th ACM SIGKDD International conference on knowledge discovery and data mining. New York: ACM, 2006:44-64.

作者贡献说明：

邹本涛：提出研究思路，负责文献调研、数据分析与论文撰写；

王曰芬：提出论文思路，指导论文修改与定稿；

余厚强：参与讨论研究思路和提出修改意见。

邹本涛，王曰芬，余厚强. 人工智能领域高产科研团队的演化研究[J]. 图书情报工作,2020,64(20):23 – 33.

Study of the Evolution Pattern of Prolific Research Teams in the Artificial Intelligence Field

Zou Bentao[1] Wang Yuefen[1,2] Yu Houqiang[1,2]

[1] School of Economics and Management, Nanjing University of Science and Technology, Nanjing 210094

[2] Jiangsu Social Public Security Technology Cooperative Innovation Center, Nanjing 210094

Abstract: [**Purpose/significance**] Teamwork has become an important form of organization for knowledge innovation today. Exploring the dynamic evolution law of scientific research teams from the perspective of dynamic networks is of great significance to promote the discovery, formation and management of scientific research teams. [**Method/process**] Taking the field of artificial intelligence as an example, this paper used the Louvain community discovery algorithm to identify research teams in the field of artificial intelligence. The extreme value distribution of the number of nodes, edges, network density, and clustering coefficients in the team cooperation network were calculated. A combination of micro and macro perspective explored the laws and characteristics of the evolution of high-yield teams in this field, aiming at revealing the intrinsic motivation of the evolution of scientific research teams. [**Result/conclusion**] From the micro perspective, the extreme value distribution of co-authored network topological indicators reveals the dynamic properties of the evolution of high-yield teams in the field of artificial intelligence; from the macro perspective, high-yield teams show evolutionary commonality in network density and network average clustering coefficients, and most teams foster more new cooperative relationships in the evolution process. In view of the evolution path of the team, the phenomenon of "small group" cooperation in high-yield teams in the field of artificial intelligence is significant, and the cooperation between "small groups" directly affects the direction of the overall team.

Keywords: artificial intelligence research team extreme value distribution dynamic evolution topological measures